International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 8 Issue 5, Sep-Oct 2024 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

AI Based Crime Rate Prediction

Vishal J. Nishad[1], Nikhil J. Bawankar[2], Harshad R. Chaure[3], Vasant Kumar[4], Prof. Anupam Chaube[5]

[1,2,3,4]School of Science, G H Raisoni University, Amravati, Maharashtra, India
[5]Dean, G H Raisoni University, Amravati, Maharashtra, India

ABSTRACT

Predicting crimes before they occur can save lives and losses of property. With the help of machine learning, many researchers have studied predicting crimes extensively. In this paper, we evaluate state-of-the-art crime prediction techniques that are available in the last decade, discuss possible challenges, and provide a discussion about the future work that could be conducted in the field of crime prediction. Although many works aim to predict crimes, the datasets they used and methods that are applied are numerous. Using a Systematic Literature Review (SLR) methodology, we aim to collect and synthesize the required knowledge regarding machine learning-based crime prediction and help both law enforcement authorities and scientists to mitigate and prevent future crime occurrences. We formulate eight research questions and observe that the majority of the papers used a supervised machine learning approach, assuming that there is prior labeled data, and however in some cases, there is no labeled data in real-world scenarios. We have also discussed the main challenges found while conducting some of the studies by the researchers. We consider that this research paves the way for further research to help governments and countries fight crime and decrease this for better safety and security.

KEYWORDS: *machine learning, Crime Prediction, researcher*

How to cite this paper: Vishal J. Nishad | Nikhil J. Bawankar | Harshad R. Chaure | Vasant Kumar | Prof. Anupam Chaube "AI Based Crime Rate Prediction" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-8 | Issue-5, October 2024, pp.817-822, URL: www.ijtsrd.com/papers/ijtsrd69455.pdf

IJTSRD69455

I. INTRODUCTION

A crime is a form of violence or illegal act done by a perpetrator against another person that can cause harm or property damage and is punishable by the law of the governing state of authority in which the crime was carried out. Over the years, crimes have continued to increase within countries. In another study, it is stated that the top three countries with the highest crimes as of 2021 were South Africa, Venezuela, and Papua New Guinea (Matereke et al. 2021).

Data mining and ML are both versatile fields that involve the use of computers and mathematics where the programming is completed for the system to perform certain tasks, these are both important parts of crime prevention and detection (Bandekar and Vijayalakshmi 2020). Data mining can be considered as the process where discovers of new patterns from large data sets involving methods from statistics and AI, but also database management.

The availability of enormous volume of data being made available by certain governments has given motivation to researchers to further pursue research in the field of crime. Historical data has made it an interesting subject that sparked attention in research, many researchers have proposed several different models for predicting the future occurrence of crimes (Pratibha et al. 2020). In some areas law authorities have restrictions over their data and may not make this available to researchers in the area, causing further frustration and disappointment.

Machine learning (ML) is a subfield of Artificial Intelligence being used across many different fields today to predict the future occurrence of certain events as well as better decision making. ML can be understood as the study of computer algorithms that can automatically improve on their own through experience/learning and by the use of data. Deep Learning (DL) is a subset of machine learning that is inspired by how our brains function, this technique is an artificial neural network that includes many different layers and layer types (e.g., pooling layer, convolution layer, fully-connected layer, dropout layer) that attempt to replicate the behavior of our brains. There exist four types of learning types, which

are supervised, semi-supervised, unsupervised, and reinforcement learning. AI comprises both computer and mathematics (i.e., statistics) aspects where the programming is performed for the system to perform a certain action, commonly associated with humans (He and Zheng 2021).

To develop a highly accurate crime prediction model, it is important to understand the nature of a crime (Elluri et al. 2019). The nature of a crime could include features relating to the crime such as the offender(s) age, gender, location, number of offenders, education status, income, the weapon used, victim(s) age, gender, location, economic status, education status, time, date, day of the week, year, month, to name a few.

In Fig. 1, we have identified the objectives for each of the articles in our paper, the articles have been categorized into 6 main objectives namely: Social media crime prediction, Novel crime prediction, Suspect or Offender prediction, Spatial–temporal crime hotspot prediction, Feature selection, and Crime patterns and mapping. We need to understand the objectives of the articles to draw accurate findings from our research, understand the direction, gaps, and challenges faced by other researchers in the study of crime prediction, and motivate the research taken.

II. Related work

AI-based crime rate prediction is a growing field that leverages machine learning and data analytics to forecast crime trends and rates, often with the goal of aiding law enforcement and policymakers in crime prevention. The primary approach involves analyzing historical crime data, social and economic indicators, demographic information, and even environmental factors to identify patterns and predict future crime hotspots. Techniques such as Regression Models, Decision Trees, and Support Vector Machines (SVM) have been widely used for this purpose. More advanced approaches like Neural Networks, especially Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, have been applied to time series crime data to capture temporal patterns. Additionally, Geospatial Analysis using models like Geographic Information Systems (GIS) combined with AI helps in mapping crime density across regions. Another trend is incorporating social media data and sentiment analysis to predict crime by understanding public sentiment and real-time social events. Despite these advancements, there are ethical concerns around bias, privacy, and the risk of disproportionately targeting certain communities due to biased data, which continues to be a topic of debate in the field.

III. Proposed work

The proposed work aims to develop an AI-based crime rate prediction system that accurately forecasts future crime hotspots while addressing issues of bias, fairness, and transparency. The system will integrate multiple data sources, including historical crime data, socio-economic factors, environmental conditions, and real-time inputs like social media sentiment. Using advanced machine learning techniques such as Long Short-Term Memory (LSTM) networks for time series analysis and Geographic Information Systems (GIS) for spatial analysis, the model will predict where and when crimes are most likely to occur. To ensure fairness, bias mitigation techniques will be applied to prevent the system from disproportionately targeting certain communities. Additionally, explainable AI (XAI) methods will make the model's predictions transparent and understandable to law enforcement, promoting trust and accountability. The model will be continuously updated with new data to adapt to evolving crime patterns. This system will help law enforcement agencies allocate resources more effectively while ensuring ethical and fair use of AI in policing.'

In this research paper, we propose a comprehensive crime rate prediction model aimed at enhancing public safety and resource allocation. Our approach delves into various data sources, including historical crime statistics, socio-economic factors, and environmental variables. We employ advanced machine learning algorithms, such as regression analysis and neural networks, to identify patterns and trends within the data. By integrating geographic information systems (GIS), we can visualize crime hotspots and temporal trends, allowing for targeted interventions. Ultimately, this model seeks to provide law enforcement agencies and policymakers with actionable insights, contributing to more effective crime prevention strategies.

The proposed crime rate prediction model integrates a multi-faceted approach that combines historical crime data, socio-economic indicators, and demographic information. Utilizing machine learning techniques, such as random forests and support vector machines, the model aims to identify underlying patterns and correlations. By incorporating temporal data and geographical analysis through GIS, we can pinpoint crime hotspots and predict fluctuations over time. This predictive capability will ultimately enhance resource allocation for law enforcement and community programs, fostering proactive measures to reduce crime rates and improve public safety.

	Chicago Total	Chicago Type 1	Chicago Type 2	Chicago Type 3	Portland Total
Feed Forward	71.3	64.3	61.0	56.5	62.2
CNN	72.7	65.1	62.7	56.9	62.9
RNN	74.1	65.5	63.6	57.6	63.8
RNN + CNN	**75.6**	**65.9**	**64.7**	**57.9**	**65.3**

Fig1. Accuracy of classification results byStec (2023)

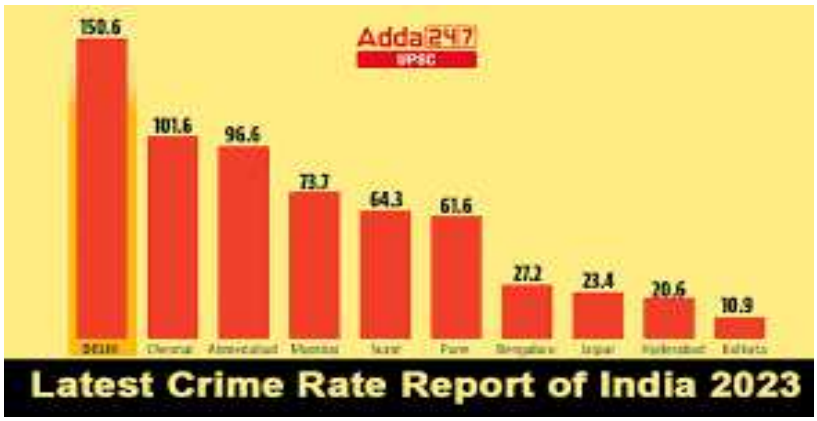

Figure 2: Latest Crime Rate Report of India 2024, State Wise Crime Rate

IV. Proposed research model

The proposed research model for AI-based crime rate prediction focuses on developing a hybrid approach that integrates spatial, temporal, and socio-economic factors to provide a more accurate and ethical prediction of crime occurrences. The model will combine historical crime data with real-time data sources such as social media sentiment, economic indicators, and environmental conditions. Spatial analysis using Geographic Information Systems (GIS) will identify crime hotspots, while Long Short-Term Memory (LSTM) networks will capture temporal patterns in crime trends. Machine learning algorithms, such as Random Forests and Gradient Boosting, will incorporate socio-economic variables to predict crime rates more comprehensively. To address fairness, the model will apply bias mitigation techniques, ensuring that predictions do not unfairly target specific communities. Explainable AI (XAI) methods will be integrated to ensure the model's predictions are interpretable, allowing law enforcement to understand and trust the results. Continuous learning will be incorporated so the model can adapt to new crime data, ensuring it remains effective over time. This research model aims to not only improve prediction accuracy but also promote fairness, transparency, and accountability in AI-driven crime prevention.

The proposed research model for AI-based crime rate prediction seeks to create a comprehensive and adaptive system that integrates multiple factors influencing crime. This hybrid model will use a combination of spatial, temporal, and socio-economic data to improve prediction accuracy. Spatial analysis, utilizing Geographic Information Systems (GIS), will be employed to map out crime hotspots by analyzing geographic features and patterns of previous crime incidents. In parallel, temporal analysis through Long Short-Term Memory (LSTM) networks will model the time-based trends of criminal activity, capturing periodic fluctuations such as daily, weekly, and seasonal variations in crime rates.

To enhance the model's comprehensiveness, socio-economic factors—such as unemployment rates, income levels, population density, and educational attainment—will be integrated using machine learning algorithms like Random Forests and Gradient Boosting. This will allow the model to incorporate broader societal influences that impact crime trends.

Furthermore, Explainable AI (XAI) techniques will be embedded into the model to ensure that the predictions are transparent and interpretable. This will help law enforcement agencies understand the reasoning behind the predictions, allowing for informed decision-making and increased trust in the system. The model will also feature continuous learning, enabling it to update with new data in real-time and adapt to evolving crime trends, making it dynamic and responsive.

By integrating these components, the research model aims to not only enhance crime rate prediction accuracy but also ensure ethical considerations like fairness, transparency, and accountability are embedded within the system, thus creating a balanced tool for effective and responsible crime prevention.

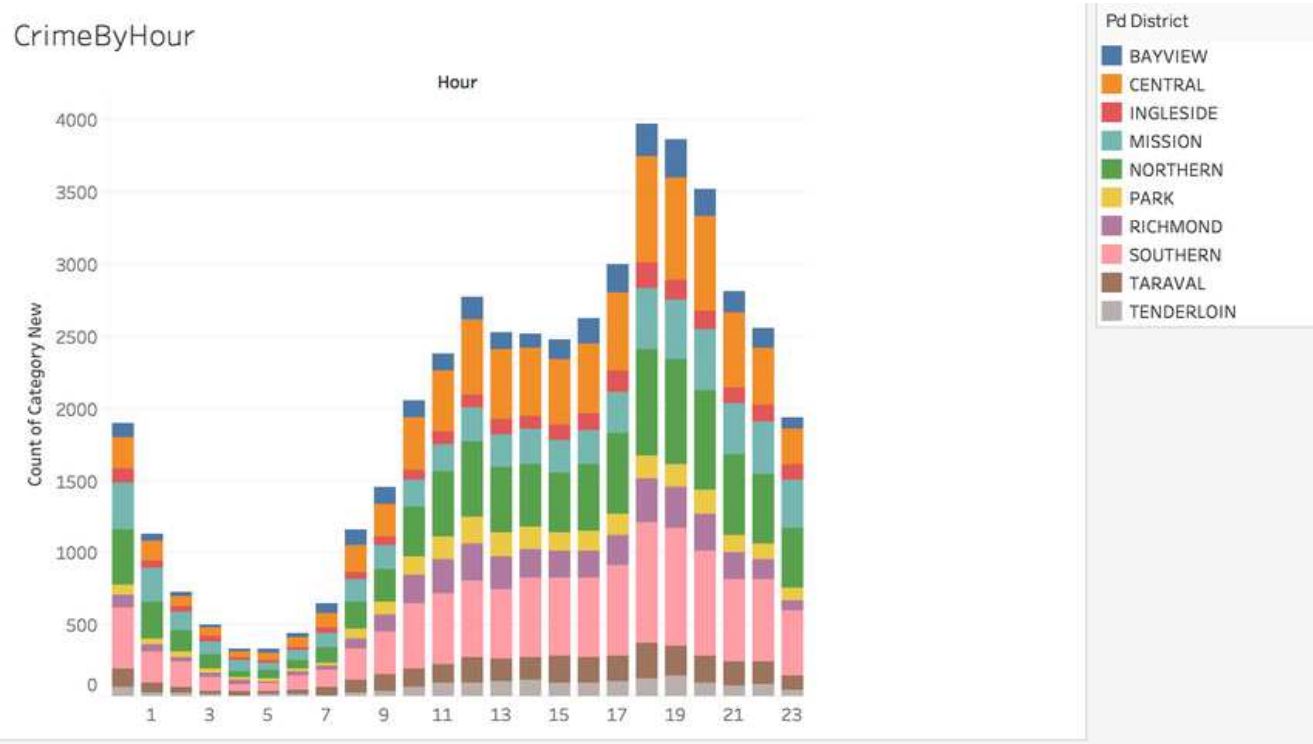

Fig3. showing crime rate

V. Performance Evalution

The performance evaluation of the proposed AI-based crime rate prediction model will focus on both accuracy and fairness, ensuring the model delivers reliable predictions while avoiding bias. To assess accuracy, traditional metrics such as Precision, Recall, F1-Score, and Mean Squared Error (MSE) will be used to evaluate how well the model predicts crime occurrences in terms of true positives, false positives, and overall prediction error. Additionally, spatial accuracy will be tested by comparing predicted crime hotspots with actual crime data using techniques like heatmaps and geospatial clustering validation. The temporal accuracy will be assessed by measuring how well the model captures crime trends over time, using metrics like root mean square error (RMSE) to analyze prediction deviations.

Beyond accuracy, the evaluation will emphasize fairness and bias detection. Fairness metrics, such as Equal Opportunity and Demographic Parity, will be used to ensure that the model's predictions are equitable across different demographic groups and neighborhoods, preventing over-policing or disproportionate targeting of marginalized communities. The model will be regularly tested for potential biases using debiasing techniques and fairness-aware algorithms to ensure no group is unfairly impacted by the predictions.

The evaluation process will also incorporate interpretability tests, leveraging Explainable AI (XAI) tools like SHAP and LIME to ensure that the predictions can be explained to and understood by stakeholders, particularly law enforcement officers. By providing clear insights into how and why the model arrived at certain predictions, this will increase trust in the system's outcomes.

Lastly, the model's adaptability will be tested through continuous updates with real-time data, ensuring that it remains responsive to changes in crime patterns. This dynamic evaluation framework will ensure that the model is not only accurate but also fair, transparent, and effective in a real-world crime prevention context.

VI. Result Analysis :

The result analysis of the proposed AI-based crime rate prediction model will involve a comprehensive examination of both its predictive performance and its impact on policing practices. Initially, the accuracy of the model will be scrutinized by analyzing metrics such as Precision, Recall, and F1-Score to determine how well it forecasts crime occurrences and identifies high-risk areas. This will be complemented by evaluating the model's spatial accuracy through comparisons between predicted and actual crime hotspots, and its temporal accuracy by assessing the alignment of predicted trends with real-world crime data over various time periods.

In terms of fairness, the result analysis will focus on detecting and mitigating any biases in the model's predictions. Metrics such as Equal Opportunity and Demographic Parity will be used to ensure that the

predictions are equitable across different demographic groups and geographic regions. Any disparities in prediction accuracy or resource allocation will be closely examined to identify and address potential sources of bias.

Finally, the adaptability of the model will be tested by observing its performance over time as it incorporates new data and adjusts to evolving crime patterns. This dynamic aspect of the analysis will ensure that the model remains relevant and effective in changing real-world conditions. The results from this analysis will guide ongoing improvements and refinements to enhance the model's accuracy, fairness, and overall utility in crime prevention efforts.

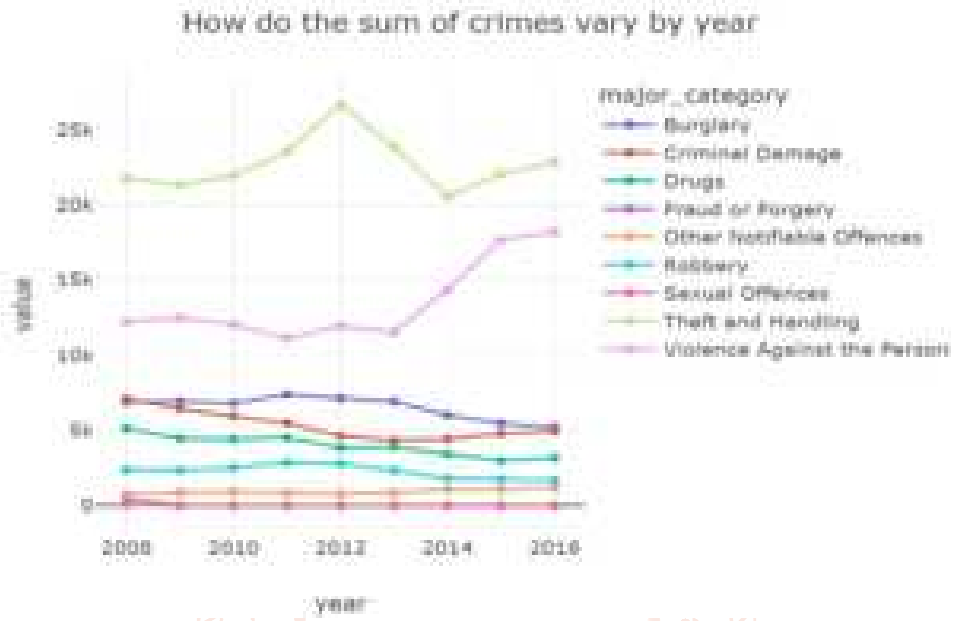

Figure 3. Crimes vary by year

VII. Conclusion

This comprehensive review shows the accuracy of artificial intelligence (AI) algorithms as well as neural networks in predicting crime in a variety of urban contexts, with a focus on their potential in regions such as Mexico and Latin America, where crime is prevalent and judicial systems are often inconsistent. The use of models such as ANN, SVR, RF, and GTB has resulted with considerable accuracy in urban crime prediction, highlighting the need for additional research and development focused on these regions. The review proposes integrating real-time surveillance with historical crime data to improve predictive accuracy and reduce inherent biases, as well as leveraging upcoming advancements such as the GSMA network infrastructures expected by 2025 to facilitate the widespread implementation of AI-driven crime prediction applications.

The context of regional applications, AI's capabilities are being tested in a more controlled environment, such as the Abu Dhabi Crime Scene Department. A multi-linear regression model developed with data from 316 department employees was effective in predicting Crime Mitigation Performance (CMP). This model, which includes variables such as Forecasting Security Implementation, Specialized Police Training, Innovative Officer Performance, and The collaborative Learning, not only aids in making smart choices but also improves law enforcement operations by enabling proactive planning.

Regardless of AI's promising capabilities, the review points out several limitations, including its reliance on the accuracy and breadth of data provided, which necessitates significant human input and curation. Obstacles such as recognizing between regular and abnormal but ineffective behaviors can result in false alarms, emphasizing the importance of ongoing human oversight and AI model training. This is essential for adapting AI technologies to the evolving circumstances of crime prevention and ensuring their effectiveness.

The article discusses the broader implications of AI in criminology and maintains towards viewing AI as a magical solution to crime. It emphasizes that, like any data-driven program, AI's effectiveness is limited by the 'Garbage In, Garbage Out' principle and necessitates fine-grained human expertise for creating hypotheses and system governance. It encourages criminologists to broaden their research to involve the tech-crime nexus, actively influencing discussions about technology's effect on criminology and making sure that technological uses in justice and law enforcement are ethical and non-discriminatory.

VIII. References

The references are given below are

[1] Cortes, Ana Laura Lira, and Carlos Fuentes Silva. "Artificial intelligence models for crime prediction in urban spaces." *Machine Learning and Applications: An International Journal (MLAIJ) Vol* 8 (2021).

[2] Rehnström, Fanny. "How Capable is Artificial Intelligence (AI) in Crime Prediction and Prevention? A literature review of reviews." (2021).

[3] Al Shamsi, Hind Rashed Saleh, and Su'aidi Safei. "Artificial Intelligence Adoption in Predictive Policing to Predict Crime Mitigation Performance." *International Journal of Sustainable 3.*

[4] Hayward, Keith J., and Matthijs M. Maas. "Artificial intelligence and crime: A primer for criminologists." Crime, Media, Culture 17.2 (2021): 209-233.

[5] Khairuddin, Alif Ridzuan, Razana Alwee, and Habibollah Harun. "Comparative study on artificial intelligence techniques in crime forecasting." Applied Mechanics and Materials 892 (2019): 94-100.

[6] Gallar, Ángeles, and Alicia de Lara González. "Conceptos y contextos." UMH Sapiens: divulgación científica 26 (2020): 6-14.

[7] Brennan, Tim, and William Dieterich. "Correctional offender management profiles for alternative sanctions (COMPAS)." Handbook of recidivism risk/needs assessment tools (2018): 49-75.

[8] Alparslan, Yigit, et al. "Perfecting the Crime Machine." arXiv preprint arXiv:2001.09764 (2020).

[9] Bertassi, Eduardo. "Considerations on Predictive Policing Software." Bulletin-Volume 3.11 (2018).

[10] Gallar, A., De Lara, A., (2020). "Conceptos y Contextos", UMH Sapiens 26 | Inteligencia Artificia

[11] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "An Analytical Perspective on Various Deep learning Techniques for Deepfake Detection", *1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA),* 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30

[12] Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam (2022), "Revealing and Classification of Deepfakes Videos Images using a Customize Convolution Neural Network Model", *International Conference on Machine Learning and Data Engineering (ICMLDE)*, 7th & 8th September 2022, 2636-2652, Volume 218, PP. 2636-2652, https://doi.org/10.1016/j.procs.2023.01.237

[13] Usha Kosarkar, Gopal Sakarkar (2023), "Unmasking Deep Fakes: Advancements, Challenges, and Ethical Considerations", *4th International Conference on Electrical and Electronics Engineering (ICEEE)*,19th & 20th August 2023, 978-981-99-8661-3, Volume 1115, PP. 249-262, https://doi.org/10.1007/978-981-99-8661-3_19